Nucleic Acids Research, 1995, Vol. 23, No. 3 341–350

Divalent metal ion binding to a conserved wobble pair defining the upstream site of cleavage of group I self-splicing introns

Frédéric H.-T. Allain and Gabriele Varani*

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Received November 18, 1994; Accepted December 21, 1994

ABSTRACT

The upstream site of cleavage of all group I self-splicing introns is identified by an absolutely conserved U•G base pair. Although a wobble C•A pair can substitute the U•G pair, all other combinations of nucleotides at this position abolish splicing, suggesting that it is an unusual RNA structure, rather than sequence, that is recognized by the catalytic intron core. RNA enzymes are metalloenzymes, and divalent metal ion binding may be an important requirement for splice site recognition and catalysis. The paramagnetic broadening of NMR resonances upon manganese binding at specific sites was used to probe the interaction between divalent metal ions and an oligonucleotide model of a group I intron ribozyme substrate. Unlike previous studies in which only imino proton resonances were monitored, we have used isotopically labelled RNA and a set of complete spectral assignments to identify the location of the divalent metal binding site with much greater detail than previously possible. Two independent metal binding sites were identified for this oligonucleotide. A first metal binding site is located in the major groove of the three consecutive G•C base pairs at the end of double helical stem. A second site is found in the major groove of the RNA double helix in the vicinity of the U•G base pair. These results suggest that metal ion coordination (or a metal bridge) and tertiary interactions identified biochemically, may be used by group I intron ribozymes for substrate recognition.

INTRODUCTION

RNA enzymes are metalloenzymes: divalent metal ions (Mg^{2+}, Mn^{2+}) are required for catalysis and also for folding (1). The functional and structural role of metal ions in RNA catalysis has only recently begun to be addressed (2–4). Because of their critical role in both catalysis and RNA folding, it is important to investigate the structural and thermodynamic properties of the interactions between divalent metal ions and RNA enzymes.

Group I self-splicing introns are one of the best characterized class of RNA enzymes (5). The secondary structure of these enzymes has been established from phylogeny and chemical and enzymatic mapping, and models have been proposed for the tertiary folding of the group I intron catalytic core (6). Intron excision is accomplished by two successive transesterification reactions in the presence of Mg^{2+} or Mn^{2+}, resulting in the excision of the non-coding region and ligation of the flanking exons. The upstream site of cleavage is located within a conserved double helical element ('P1 helix') and is defined by an absolutely conserved U•G base pair and by 4–5 2′-OH groups in the nucleotides near the splice site (7–9). The wobble pair is critical for substrate recognition: mutation of the U•G base pair to a perfect Watson–Crick pair abolishes ribozyme activity. Of all 16 possible base pair combinations, only U•G and C•A wobble pairs are tolerated (10). RNA structure, rather than sequence, determines the cleavage specificity of group I self-splicing introns.

A number of investigations have been dedicated to clarifying the roles of metal ions in group I introns. Based on the different ability of Mg^{2+} and Mn^{2+} to coordinate sulfur, it was suggested that magnesium and manganese ions interact with a non-bridging phosphate (4) and with the 3′ oxygen of the uracil at the upstream site of cleavage (5′ splice site) (3). The details of the metal binding sites and the nature of the interactions are not known, although models have been proposed based on known structures of protein metalloenzymes (11,12).

If a paramagnetic ion like manganese is located in the vicinity of an NMR-active nucleus (within a radius of approximately 1 nm), the relaxation rate of the nuclear spin increases, resulting in broadening of the corresponding NMR resonance (13). The increase in the rate of relaxation is proportional to r^{-6}, where r is the distance between the observed nucleus and the paramagnetic ion, and to the degree of occupancy of the metal binding site. At high concentrations of paramagnetic impurities, all resonances are broadened beyond detection. If the divalent paramagnetic impurity is preferentially localized at specific sites on the RNA, resonances corresponding to nuclei located in the vicinity of the metal binding site are selectively broadened at very low concentration of Mn^{2+}. The observation of selectively broadened resonances in the RNA spectrum, and the distance dependence of

* To whom correspondence should be addressed

the increase in relaxation rate, allow the localization of the metal binding sites.

Using the paramagnetic broadening of NMR resonances induced by small amounts of Mn^{2+} ions, it was recently shown that a G•U base pair defines a divalent ion binding site in a tRNA acceptor stem (14,15). This base pair is a critical determinant of tRNA recognition by its corresponding amino acyl tRNA synthetase. By analogy, it is possible that one feature of the conserved U•G base pair at the 5′-splice site of all group I introns is its ability to bind divalent metal ions in a similar manner to that observed in the tRNA acceptor stem. In order to test this hypothesis, and to refine current methodologies to identify metal binding to RNA oligonucleotides in solution using NMR, we investigated the interaction of Mn^{2+} ions with an oligonucleotide representing the substrate for group I intron catalysis. In order to stabilize the RNA secondary structure, the phylogenetically variable loop at the top of the P1 stem was substituted with an exceptionally stable tetraloop sequence (16). Unlike previous studies which were limited to the observation of paramagnetic broadening in the well-resolved imino proton region of the RNA NMR spectrum, the entire NMR spectrum, and consequently the RNA structure, was probed by using several multidimensional heteronuclear experiments with isotopically labeled RNA. Two independent metal binding sites were identified for this oligonucleotide. The first is located in the major groove of the RNA double helix in the vicinity of the U•G base pair. A second metal binding site is located in the major groove of the three consecutive G•C base pairs at the end of double helical stem.

MATERIALS AND METHODS

RNA preparation and purification

The oligonucleotide 5′GGGAUAACUUCGGUUGUCCC and several mutant sequences were synthetised using T7 RNA polymerase and synthetic DNA templates containing a double stranded 17 bp T7 promoter region and a single stranded template (17). The RNA was purified by 20% polyacrylamide gel electrophoresis, electroeluted (Schleier and Schuell), ethanol precipitated and resuspended in a minimal volume. The terminal 5′-phosphates were removed by dephosphorylation with alkaline phosphatase (Boehringer). The phosphatase was then removed by phenol–chloroform extraction, and the RNA was desalted by size-exclusion sephadex G-15 chromatography before extensive dialysis against the final NMR buffer (50 mM NaCl, 10 mM sodium phosphate, pH 5.5). RNA concentrations were determined from the UV absorbance at 260 nm (18). Four different mutants (see text) and the wild-type P1 helix were prepared by an identical procedure, using template DNAs of the appropriate sequence.

Preparation of isotopically labelled RNA

Isotopically labelled RNA was prepared in the same way as the unlabelled RNA, but the T7 polymerase transcription reaction was done with ^{13}C and/or ^{15}N uniformly labelled NTPs (19–21). Briefly, ^{15}N and ^{13}C labelled NTPs were prepared by extracting ribosomal RNA from *E.coli* grown on M9 minimal media containing 99% pure ^{15}N ammonium sulphate as the sole nitrogen source and 99% ^{13}C glucose as the sole source of carbon. The rRNA was hydrolysed to the individual nucleotides with nuclease P1 (Sigma); completion of the reaction was verified by thin layer chromatography on PEI-cellulose plates (Merck). The mononucleotides were then rephosphorylated enzymatically using phosphoenol pyruvate (PEP) as the phosphate source. The completion of the reaction was monitored by HPLC using a Vydac-3 ion-exchange analytical column. The NTPs were ethanol precipitated, resuspended in a minimal volume, phenol–chroroform extracted and ethanol precipitated again. NTPs were desalted to improve transcriptional efficiency on a G10 (Sephadex) size exclusion column running in a 1 mM Tris–HCl, pH 8.1 buffer. Approximately 0.5 ml of 1 mM P1 helix RNA were obtained from a total reaction volume of 20 ml (4 mM in each NTP). This represents a 4% NTP incorporation for the final, fully purified product.

Manganese titration

In order to observe paramagnetic broadening, the RNA samples were titrated with increasing amounts of $MnCl_2$ (4–200 μM, the RNA concentration is ~1 mM). Different amounts of Mn^{2+} were necessary to obtain comparable paramagnetic broadening of the RNA resonances in different preparation of the same oligonucleotide, presumably because of the presence of varying concentrations of the chelator EDTA left over from the purification procedure.

NMR spectroscopy

NMR spectra were obtained on Bruker AMX500 or DMX600 spectrometers equipped with triple resonance probes and operating (^1H frequencies) at 500 and 600 MHz, respectively. The spectra were referenced to TSP (^1H and ^{13}C) and phosphoric acid (^{31}P), either directly or by comparison with published values (22), or to external NH_3 (^{15}N, D. Live, personal communication). Data were processed using the program FELIX 2.10 (Biosym, San Diego). For all 2-dimensional spectra, 1024 complex points were acquired in t_2 and 512 in t_1, employing the TPPI scheme (23). The processed spectra were zero-filled to a final 2k × 2k (real) data size after apodisation with 40–60° shifted sine bell functions.

Samples used in multidimensional spectra recorded in the presence of manganese were prepared by careful titration of Mn^{2+} into the sample to reach broadening comparable to that observed in Figure 2b. Samples in D_2O were prepared by the same titration procedure, lyophilized and resuspended in 100% D_2O (Aldrich).

The 1-D proton NMR spectra in H_2O were recorded at 0.4–1 mM RNA concentrations, at a temperature of 275 K using the jump–return scheme for water supression (24). The residual water signal was further reduced by digital deconvolution. Typically, 128 scans of 1024 complex points were collected, with a sweep width of 12 500 Hz. The 2-D NOESY spectra in H_2O were recorded at 1 mM RNA concentration and at 300 ms mixing time, under the same conditions as the 1-D spectra and using the jump–return water suppression scheme on the final ('read') pulse of the sequence.

The 2-D NOESY (60, 100, 120, 200 and 300 ms mixing times) and DIPSI-based (25) TOCSY (60 ms mixing time) spectra in D_2O were recorded at 1 mM RNA concentration, at a temperature of 293 K using very low power pre-saturation on the residual HDO resonance. The spectral width was 8064 Hz in both dimensions, and the total acquisition times ≈20 h.

A series of ^{31}P–^1H correlation spectra were acquired using the AMX-500 spectrometer equipped with an inverse probe. Two-

dimensional Het-Cor (26), Hetero-Tocsy (27,28), HSQC (29) and HSQC-NOESY (30) (tuning delay was 25 ms for both experiments, the mixing time was 200 ms for the NOESY portion of the experiment) spectra were acquired in the presence and absence of Mn^{2+} with spectral widths of 2500 Hz (^1H) and 1000 Hz (^{31}P). 128 FIDs of 512 real points were collected ($t_2^{max} \approx 160$ ms, $t_1^{max} \approx 64$ ms) with total acquisition times of ≈ 20 h.

The 2-D ^1H–^{15}N (HMQC) spectra (31) were recorded in the presence and absence of Mn^{2+} with delays and spectrometer frequencies optimized for detection of imino or amino RNA resonances (32) on a 1 mM ^{15}N labelled sample, at 275 K. Spectral widths were 12 500 Hz (^1H) and 2500 Hz (^{15}N). The 2-D HSQC ^1H–^{15}N spectra were recorded in D_2O, tuned to observe long range couplings and the refocusing delay before acquisition was eliminated (33). Acquisition times were 8–12 h (imino-optimized spectra) or ≈ 20 h for spectra recorded in D_2O or optimized for amino proton detection. These spectra were recorded at 293 K with ^1H spectral width of 8064 Hz and ^{15}N spectral width of 6000 Hz. A ^1H-decoupled ^{15}N 1-D spectrum was recorded in several hours on the DMX 600 spectrometer at 300 K with a spectral width of 40 000 Hz.

The 2-D ^1H–^{13}C HSQC spectra were recorded at 300 K in the presence and absence of Mn^{2+} with spectral widths of 8064 Hz (^1H) and 20 000 or 7000 Hz (^{13}C), with ^{15}N–^{13}C-decoupling during acquisition using GARP (34). 2-D HCCH–TOCSY experiments (35) were recorded with spectral widths of 8064 and 2500 Hz (first and second dimension, respectively), and with editing of the methine/methylene resonances (36). 512 FID's of 512 complex points were collected for a total acquisition time of ≈ 22 h per spectrum.

A 3-D ^{13}C-edited NOESY-HMQC spectrum (37) was recorded in ≈ 66 h at 200 ms mixing time with ^{15}N–^{13}C GARP decoupling during acquisition. The spectral widths were 8064 Hz (each ^1H-dimension) and 6000 Hz (^{13}C). 300 points were recorded in the first ^1H dimension (t_1^{max} = 18.6 ms) and 64 in the ^{13}C-dimension (t_2^{max} = 5.3 ms). The data were processed to a final size of 1024 × 128 (^{13}C) × 512 real points. Folded resonances (C8, C6, C2) had opposite phase to all other resonances (38).

Three-dimensional ^1H–^{31}P–^{13}C HCP (39) and ^{13}C-edited Het-Cor spectra (G. Varani *et al.*, submitted) spectra were recorded in 22 h each at 600 Mhz with spectral widths of 3004 Hz (^1H), 1000 Hz (^{31}P) and 4000 Hz (^{13}C). Sixty-four points were collected in each indirectly detected dimension (t_1^{max} = 32 ms, t_2^{max} = 8 ms), and the final size of the processed data was 512 × 128 (^{13}C) × 64 (^{31}P) real points.

Sequence-specific assignments

Assignments of exchangeable (imino and amino) ^1H resonances were obtained from the 2-D NOESY spectra recorded in H_2O and ^1H–^{15}N HMQC spectra. Uracil and guanine NH resonances were first distinguished from the chemical shift of the attached ^{15}N resonances in the HMQC spectrum (32). Sequence specific assignments were then obtained from the sequential connectivties observed in the NOESY spectrum. The imino resonances of the four different mutants were assigned by comparison with the wild-type molecule or, when unambiguous assignments proved impossible, by collecting 2-D NOESY spectra.

Assignment of all the non-exchangeable ^1H resonances was accomplished using well established procedures (40). Homonuclear spectra allowed assignments of all base and anomeric (H1′) resonances and most H2′ and H3′ sugar resonances. Assignments of the sugar resonances were completed using the HCCH–TOCSY spectra. ^{13}C assignments were obtained from the analysis of 2-D-HSQC and 3-D-NOESY–HMQC spectra. Base ^{15}N-resonances were assigned from HMQC spectra acquired in H_2O (imino NH and amino NH_2 resonances) or in D_2O (all others). ^{31}P-resonances were assigned using several 2- and 3-dimensional spectra. The combination of the 3-D ^{13}C-edited Het-Cor spectrum with the 3-D HCP spectrum proved very powerful. In one case (Het-Cor) the strongest cross-peaks in the spectrum generally correspond to sequential H3′-P connectivities, whereas sequential and intraresidue C4′-P and P-C4′ connectivities generate the strongest cross-peaks in the HCP spectrum. While either the C4′–H4′ (HCP) or C3′–H3′ regions (Het-Cor) are often too overlapped to allow for unambiguous assignments, the combination of these two complementary experiments allowed complete ^{31}P resonance assignments.

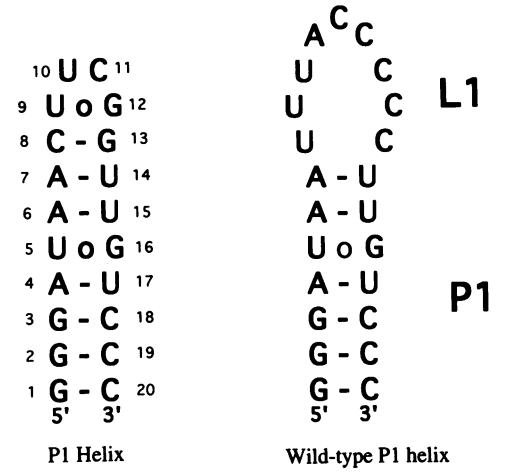

Figure 1. Sequences and secondary structures of the P1 element from the group I self-splicing intron from the large ribosomal RNA of the yeast *S.cerevisiae* (right) and of the oligonucleotide model used in this study (left). The model oligonucleotide is referred to as 'P1 helix' in the text.

RESULTS

The 3′-terminal end (NCCA) of $tRNA^{Ala}$ defines a manganese binding site in the vicinity of a G•U base pair essential for tRNA recognition and aminoacylation by its cognate synthetase (14). By analogy, we decided to examine whether a divalent ion binding site exists in the vicinity of the invariably conserved U•G base pair that defines the 5′-splice site within the P1 helix of all group I self splicing introns (10). Although base pairing is required, the precise sequence of the base pairs flanking the U•G pair is not conserved among different group I introns, nor required for self-splicing (6). The P1 helix of one of the introns of the large ribosomal RNA of the yeast *Saccharomyces cerevisiae* mitochondria (6) was chosen for this investigation for synthetic reasons. Since the isolated wild type P1 helix was not stably folded, the dispensable apical loop L1 was replaced with a stable tetraloop UUCG sequence to stabilize the RNA hairpin structure (Fig. 1); there are examples of P1–L1 elements containing stable L1 tetraloops (6). As expected, this modified hairpin is very stable and properly folded. Formation of the wobble U•G base pair

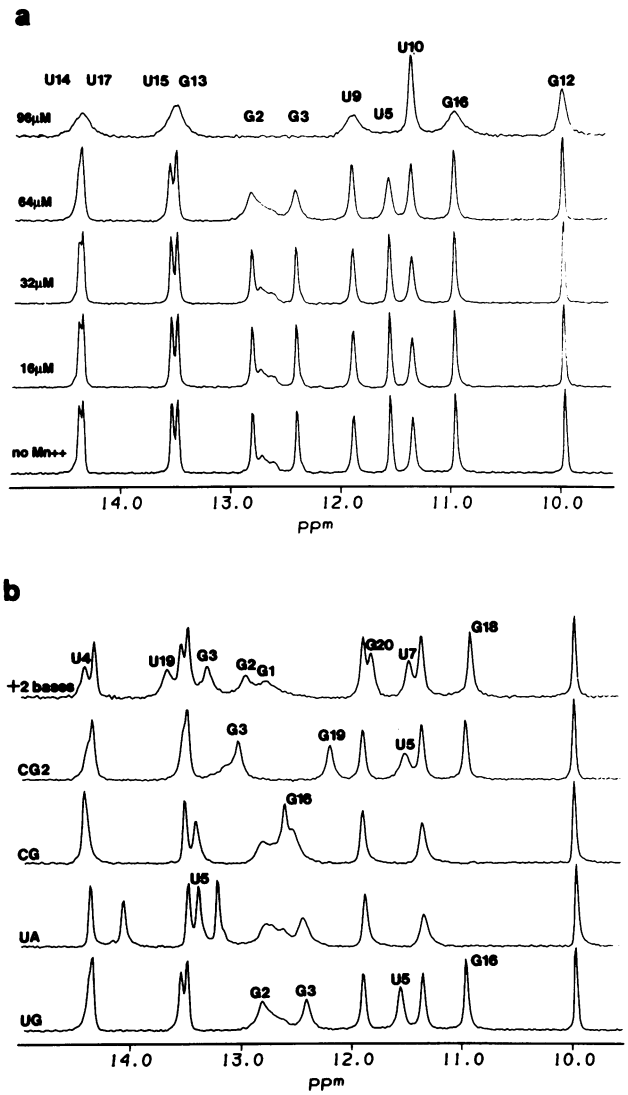

Figure 2. (a) Manganese titration of the 'P1 helix' from 1-D spectra acquired in H_2O at 275 K at increasing concentrations of Mn^{2+}, as indicated next to each spectrum. Assignments are indicated at the top of the figure. (b) Comparison of 1-D spectra of the P1 helix and four mutants at comparable stages of the Mn^{2+} titration. From the bottom, the spectra refer to the wild-type oligonucleotide and to the four mutants shown in Figure 3. Assignments of the broadened resonances are indicated for each spectrum.

within a regular A-form helix was clearly established from the assignments and the pattern of NOE interactions. The following results describe the NMR characterization of the interactions with Mn^{2+} ions of this oligonucleotide model of the P1 helix.

Manganese binding selectively broadens imino resonances at the U•G base pair and at the bottom of the helical stem

Between 10 and 15 p.p.m., the ^1H NMR spectrum of RNA presents a well resolved region where only imino resonances of guanine and uracil bases are found. The imino spectrum was assigned from a 2-D NOESY spectrum and from ^1H–^{15}N 2-D-correlation spectra optimised for detection of the imino proton resonances (see Material and Methods). When titrating the P1 helix oligonucleotide with increasing amounts of manganese

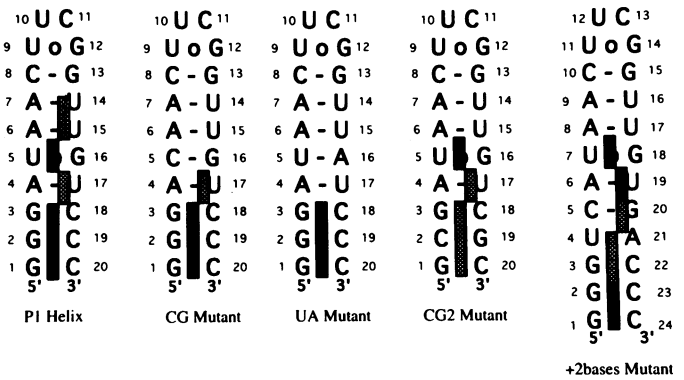

Figure 3. Summary of the results of manganese titration with the P1 helix and four mutants. Dark bars identify sites of strong paramagnetic broadening; grey bars identify sites of weaker but still significant broadening.

(from 0 to 128 μM $MnCl_2$), increased broadening of the imino resonances was observed (Fig. 2a) due to the paramagnetic effect of Mn^{2+} on the relaxation of the ^1H NMR resonances. At the highest Mn^{2+} concentration, all resonances are very broad and eventually disappear from the spectrum. This effect can be reversed by the addition of sufficient amounts of EDTA. At low Mn^{2+} concentrations, G2, G3 and U5 imino resonances are broadened while all other resonances are essentially unaffected, indicating a specific interaction. Since all RNA molecules in solution are affected by the manganese (far below stoichiometric amounts) to the same extent, observation of broadened resonances indicate that the Mn^{2+}–RNA interaction occurs in the fast exchange regime ($k_{off} > 10^3\ s^{-1}$). In turn, this implies a lower limit $K_d > 10^{-4}$ M of the binding constant, assuming the on rate is diffusion limited. A $K_d \approx 10^{-3}$ M would be comparable to that recently observed for divalent ion binding to a ribosomal RNA structural element (41). Since the U5 imino is involved in the U•G wobble base pair that is the subject of this investigation, mutant RNA sequences were used to better define the role of this wobble pair on Mn^{2+} binding.

Four different RNA mutants were transcribed (Fig. 3). In two of the mutants, the U5•G16 base pair was mutated to a Watson–Crick base pair, and in the remaining RNA mutants the end of the stem was modified. The manganese titration shown in Figure 2a was repeated for each mutant (data not shown). The comparison of different RNA sequences at comparable stages in the titration is shown in Figure 2b: at this level of broadening, only the U5 imino resonance and the two terminal G imino resonances are affected by the paramagnetic ions. When the U•G wobble pair was mutated to a U•A pair (UA mutant) or a C•G base pair (CG mutant), broadening was still observed at the end of the stem, but the imino resonances from the Watson–Crick pairs replacing the wobble pair were not affected by the presence of paramagnetic impurities. Thus, the U•G wobble pair appears necessary for manganese binding at position 5 of the stem. When the end of the stem was mutated (CG2 mutant), the effect on the wobble pair was preserved (U5 strongly broadened), whereas the resonances at the end of the stem were sharper. The two effects on the wobble pair and the terminal GGG sequence are clearly separated in the longer mutant (+2 bases mutant) where two additional base pairs were inserted between the U•G pair and the three terminal Gs. Strong broadening was observed at the U•G pair (U7 and U19) and at the bottom of the stem (G1 and G2), indicating the presence

Table 1.

Nucleotide	H8/6	C8/6	H2/5	C2/5	imino	N1	N3	PuN7	PuN9	H amino	^{15}N amino
5' G1	7.95	138.1	na	na	12.60	148.3		234.1	172.2		
G2	7.50	136.1	na	na	12.83	149.2		235.4	172.1	8.38	
G3	7.24	135.2	na	na	12.41	148.7		236.6	171.7	8.20/6.02	
A4	7.66	138.7	7.82	153.6	na	215.6	224.9	233.4	173.4	7.97/6.80	85.82
U5	7.55	139.8	5.40	103.6	11.55	147.6	160.2	na	na	na	
A6	8.31	140.2	6.65	151.9	na	214.5	224.2	232.9	173.2	7.95/6.65	83.82
A7	7.89	138.7	7.79	153.6	na	213.8	224.9	232.5	173.5	8.50/6.80	87.33
C8	7.29	139.6	5.14	96.6	na	153.5		na	na	8.50/7.11	100.7
U9	7.76	140.2	5.67	104.5	11.87	149.2	161.5	na	na	na	na
U10	8.04	144.2	5.88	104.9	11.34	146.3	160.0	na	na	na	na
C11	7.72	142.4	6.13	97.9	na	153.0		na	na	7.20/6.43	95.0
G12	7.89	142.4	na	na	9.95	144.7		234.1	174.0		
G13	8.38	138.7	na	na	13.48	149.6		235.0	172.7	8.77/6.77	75.38
U14	7.87	141.8	5.14	101.7	14.34	149.7	163.7	na	na	na	na
U15	7.98	141.8	5.71	103.3	13.54	149.1	164.4	na	na	na	na
G16	7.77	136.8	na	na	10.95	145.2		237.6	172.5	6.35	
U17	7.78	140.7	5.20	102.3	14.37	149.0	164.4	na	na	na	na
C18	7.92	141.3	5.62	96.7	na	154.0		na	na	8.46/7.09	
C19	7.81	141.0	5.50	96.9	na	153.6		na	na	8.51/7.01	
3' C20	7.69	141.1	5.51	97.4	na	154.8		na	na	8.38/7.13	

of two distinct binding sites for Mn^{2+} with comparable affinity. A fifth mutant containing a C•A wobble pair instead of the U•G wobble pair was also tested, but the absence of an imino resonance for this base pair did not allow us to draw any conclusion on the preservation or loss of Mn^{2+} binding to this mutant.

Taken together, the results summarized in Figure 3 strongly indicate that this P1 helix oligonucleotide model defines two Mn^{2+} binding sites with comparable affinity. The first is located at the end of the stem, where three consecutive Gs on the same strand are necessary, and a second site requires the wobble U•G pair. Since terminal phosphates have been removed by dephosphorylation, binding at the end of the stem cannot be simply due to electrostatic interactions with the negatively charged terminal triphosphates.

Manganese binding occurs in the major groove of the RNA helix

The above results indicate that this RNA sequence defines two metal binding sites. The precise location of those sites was investigated by identifying all nuclei that are located in the vicinity of the Mn^{2+} ions. Given the r^{-6} dependence of the paramagnetic effect, this information, together with the forthcoming high resolution structure (Allain and Varani, submitted) would allow the precise location of the Mn^{2+} with respect to the RNA bases and phosphodiester backbone. This section describes the systematic comparison of multidimensional homonuclear and heteronuclear NMR spectra acquired in the presence and absence of Mn^{2+} paramagnetic impurities at a level where only specific broadening of the U•G and the terminal GGG sequence is observed. Since complete resonance assignments are available for all ^{1}H and ^{31}P resonances and all ^{15}N and ^{13}C resonances directly attached to protons (Tables 1 and 2, see Materials and Methods), comparison of spectra obtained with and without Mn^{2+} should provide a precise location of the metal binding site. When possible, the effect of the paramagnetic ions was directly measured from the comparison of 1-D spectra with and without Mn^{2+}, but spectral overlap required the comparison of homonuclear and heteronuclear 2-D spectra. All base resonances broadened in the presence of manganese are shadowed in Table 1.

The paramagnetic effect due to specific Mn^{2+} binding on H8/H2 purine resonances and on H6/H5 pyrimidine resonances and on their attached carbons can be observed in the comparison of 2-D ^{1}H–^{13}C correlated (HSQC) spectra acquired in the presence and absence of Mn^{2+} (Fig. 4). The H8–C8 cross-peaks of the guanine G1, G2 and G3 are broadened beyond detection in presence of Mn^{2+} (Fig. 4a). A weaker broadening effect can be detected on the H8 resonance of A4, but not on the corresponding C8 resonance. All other cross-peaks are not broadened, confirming that a first binding site is near the three terminal guanine bases. A second section of the same spectra (Fig. 4b) shows the clear Mn^{2+} broadening effect on the H5–C5 cross-peaks of U5, U14, U15 and U17, while none of the other H5–C5 and H1′–C1′ cross-peaks are affected. It should be noted that the H6–C6 resonances of the corresponding bases are not broadened (Fig. 4a). Therefore, this technique can identify the location of the Mn^{2+} ions to a resolution comparable to the 2.4 Å separation between pyrimidine H6 and H5 protons.

The paramagnetic broadening of the amino protons was observable in NOESY spectra (data not shown). With the exception of the amino resonances of A7, C8, G13 and G16, all amino protons are broadened in the presence of manganese. The broadening of the base ^{15}N resonances was observed in a 1-D ^{15}N spectrum and in a 2-D ^{1}H–^{15}N correlated spectrum acquired in D_2O and optimized for detection of long range correlations to aromatic and anomeric H1′ resonances (33). Both the H8–N7 region of the HSQC spectrum (Fig. 5) and the ^{15}N 1-D data (not

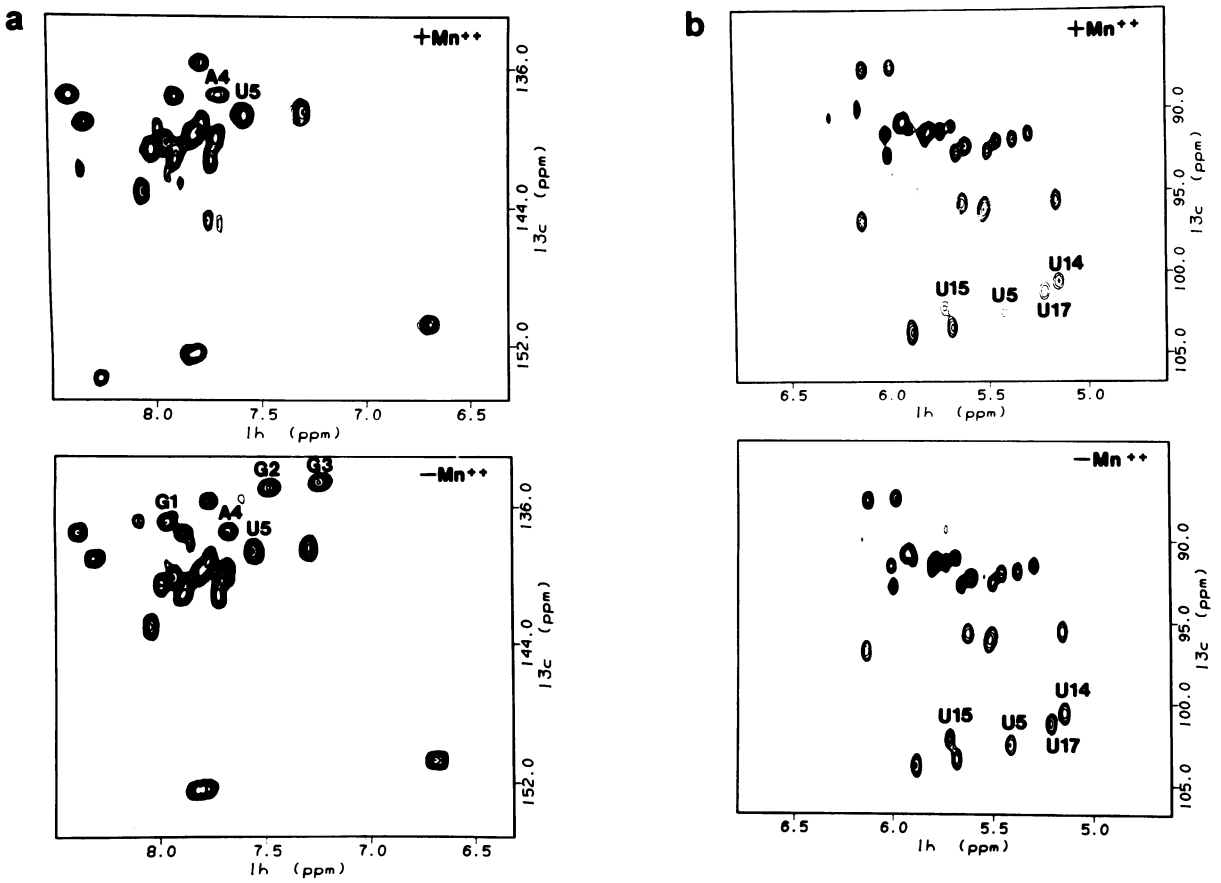

Figure 4. Comparison of ^{1}H–^{13}C correlated (HSQC) spectra of the P1 helix in the presence (top) and absence (bottom) of manganese impurities. The region shown in (**a**) contains correlations between base resonances (H2, H8 and H6) and their attached carbons. The region shown in (**b**) contains correlations between pyrimidine H5 and all H1′ resonances and their attached carbons.

shown) show clear broadening of the G1, G2, G3, A4 and G16 N7 resonances. The effect on the N9 nuclei is less clear, but the absence of H1′–N9 cross peaks for G1, G2 and G3 (it is known from the 2-D ^{1}H–^{13}C HSQC that the H1′ are not broadened) suggest that the G1, G2, G3 N9 nitrogens are in the vicinity of a manganese binding site. In contrast, A4 and G16 N9 are not affected, since the H1′–N9 cross peaks are clearly unaffected in the sample with manganese. Not surprisingly, the G1, G2, G3 H1–N1 cross peaks are broadened in the HMQC spectrum optimised for imino proton detection, but it was not possible to see a direct effect on the N1 resonance from the ^{15}N 1-D spectra, since the region where the N1 resonances are located is too crowded. It should be also noted that all adenine N1 and all guanine N3 resonances are not broadened under our experimental conditions. The results confirmed that none of the pyrimidine N1s are broadened, and therefore must be located away from the Mn^{2+} binding sites. The U5 and U17 N3 resonances are broadened, but, as in the case of the guanine N1, it is not clear whether only the H3 proton is broadened, or the corresponding N3 resonances are broad as well.

The results of the previous paragraphs demonstrate that the manganese binding sites at the U•G wobble pair and at the bottom of the stem are located in the major groove of the RNA helix. As shown in Figure 6, where resonances affected by the Mn^{2+} paramagnetic properties are shadowed, many resonances located in the major groove are in close contact with the Mn^{2+} ions, for instance the purine N7/H8 resonances of the three consecutive Gs at the end of the helical stem. On the contrary, the N3 purine resonances and adenine H2 resonances in the minor groove are not affected at all. For the U•G wobble pair, the U5 NH and H5 protons, as well as the N7 of G16, all in the major groove, are broadened by the addition of Mn^{2+}, whereas the G16 NH, NH2 and N3 resonances in the minor groove are not.

Manganese-backbone interactions are limited to G3 and U5

The paramagnetic effect of Mn^{2+} on the phosphodiester backbone is summarized in Table 2, where resonances of the ribose and phosphate moieties broadened in the presence of manganese are shadowed. The 2-D ^{1}H–^{13}C correlated spectra clearly show that none of the H1′ and C1′ resonances are broadened in the presence of manganese (Fig. 4b). Because of the spectral overlap, identification of broadened sugar resonances was obtained from 2-D HCCH–TOCSY spectra (not shown). Only the H2′ resonances of G1, G2 and G3, the H3′ resonance of U5, the H4′ of G1 and the H5′/H5″ resonances of G1, G3 and A4 were broadened.

Since the phosphates are negatively charged, one could expect that the interaction between the manganese and the RNA would take place by a direct interaction with the phosphate moiety. The

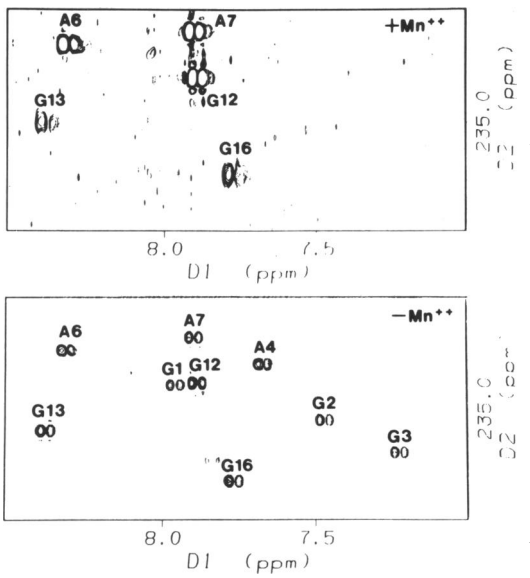

Figure 5. Comparison of the H8–N7 regions of ^{1}H–^{15}N correlated spectra (HMQC) optimized for observation of long range correlations in D_2O. The spectrum at the top was acquired in the presence of Mn^{2+}, whereas the spectrum at the bottom was acquired without Mn^{2+}.

^{31}P resonances were assigned as described in the experimental section, and the manganese broadening was observed by a comparison of 1-D ^{31}P spectra and various 2-D ^{31}P–^{1}H correlation spectra. The data demonstrated that only the G3 and U5 ^{31}P resonances are significantly broadened, indicating their close proximity to a Mn^{2+} binding site (Fig. 7). Thus, it appears that Mn^{2+} ions interact primarily with the RNA bases, at least in the case of the GGG sequence at the bottom of the stem. The broadening effect on the H3′ of U5 suggests the proximity of the Mn^{2+} ion to the bridging phosphate between U5 and A6. This may be particularly interesting, since this is where the 5′ splicing reaction takes place.

Modelling the manganese binding sites into the RNA structure

The structure of the P1 helix is shown in Figure 8a, and atoms found in the close proximity of the manganese ions are identified by their Van der Waals spheres. The structure shown in Figure 8 was generated using approximately 700 NOEs and dihedral angle constraints (F. Allain and G. Varani, submitted). This particular representation was chosen from a total of 20 converged structures with very similar energy, very low violations of NMR restraints and essentially superimposable structural features. Inspection of Figure 8a confirms that manganese binding occurs in the major groove of the double helix in the vicinity of the U•G wobble pair and at the run of three consecutive Gs at the bottom of the stem. The similarities with the metal binding site identified crystallographically at consecutive guanines in tRNA is highlighted by the close-up view of the three consecutive guanines at the 5′-end of the stem (Fig. 8b).

DISCUSSION

Structural information on metal binding to RNA comes almost exclusively from crystallographic studies (2,42,43). One Mn^{2+} binding site was found in the D-loop of $tRNA^{phe}$: the metal is directly coordinated to an N7 imidazole nitrogen (15,16). Four Zn^{2+} ions were also found to be coordinated to N7 of various guanines in crystallographic studies of $tRNA^{ala}$ (44). One site coincides with the Mn^{2+} site previously cited, and two others are localized in the major groove of adjacent purines in the T-stem and anticodon stem. Solution studies of divalent metal ions binding to tRNA and tRNA-derived oligonucleotide models (13–15) have been limited to the NMR analysis of paramagnetic broadening of imino proton resonances, and therefore somewhat limited in their structural information content.

In this study, we have used NMR spectroscopy to identify two different manganese binding sites within a stem–loop structure representing the P1 helix of group I intron self-cleaving

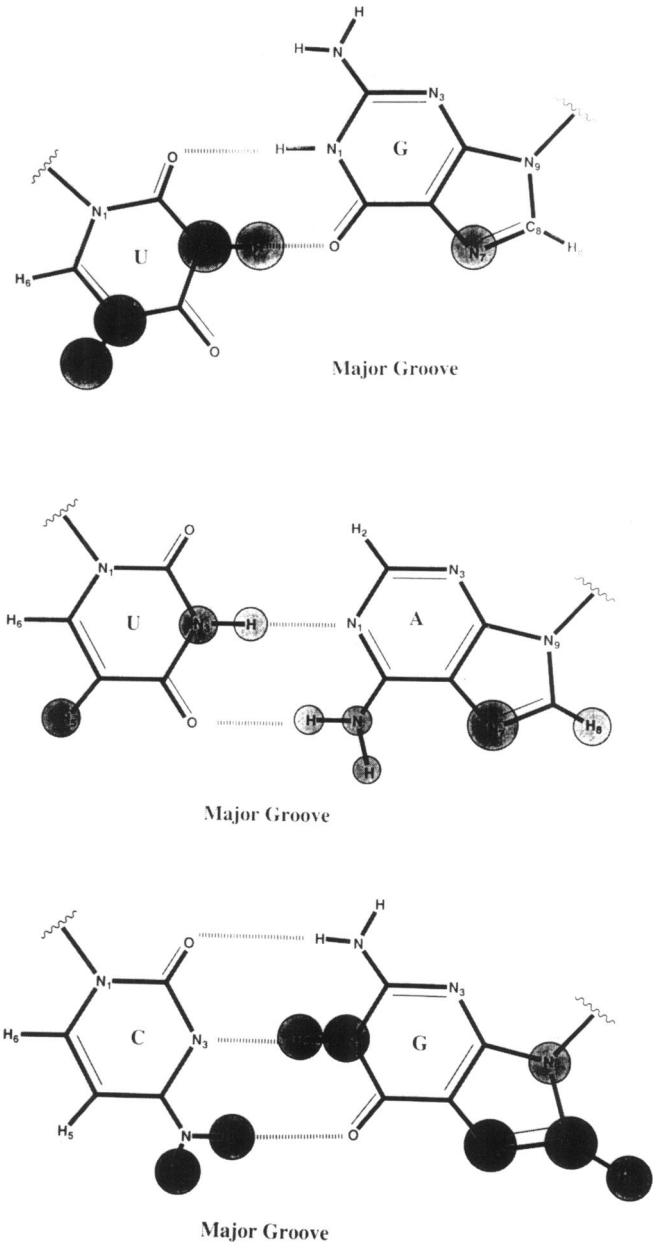

Figure 6. All base nuclei in close proximity of the metal binding site are located in the major groove (shaded circles), clearly defining this part of the double helical structure as the region where metal binding occurs.

Table 2.

Nucleotide	P	H1'	H2'	H3'	H4'	H5'/H5''	C2/5	C1'	C2'	C3'	C4'	C5'
5' G1	na	5.69	4.83	4.53	4.29	3.98/3.89	na	92.1	74.8	73.5	83.9	62.0
G2	-4.04	5.91	4.70	4.65	4.50	4.57/4.20	na	92.0	75.1	72.2	81.5	64.3
G3	-3.74	5.79	4.72	4.54	4.65	4.50/4.12	na	92.4	75.1	72.5	81.5	65.0
A4	-4.11	6.01	4.71	4.46	4.53	4.57/4.14	153.6	92.7	75.0	72.1	81.5	64.3
U5	-4.06	5.29	4.20	4.55	4.43	4.52/4.09	103.6	92.6	75.1	72.4	82.0	63.6
A6	-4.30	5.92	4.61	4.80	4.50	4.57/4.20	151.9	92.0	75.2	72.7	81.5	64.8
A7	-4.22	5.94	4.51	4.59	4.52	4.57/4.20	153.6	92.0	75.1	72.2	81.7	64.8
C8	-4.30	5.38	4.32	4.18	4.40	4.46/4.03	96.6	93.0	74.9	71.2	81.4	64.0
U9	-4.43	5.64	3.77	4.54	4.36	4.49/4.09	104.5	93.8	75.4	72.5	81.8	64.0
U10	-3.55	6.13	4.67	4.04	4.50	4.26/4.06	104.9	88.7	74.0	77.1	86.5	67.2
C11	-5.17	5.97	4.12	4.50	3.82	3.63/2.73	97.9	88.5	77.2	79.7	83.9	66.7
G12	-5.05	5.99	4.83	5.68	4.43	4.45/4.21	na	93.9	76.9	75.3	82.8	68.4
G13	-2.44	4.41	4.46	4.26	4.42	4.52/4.30	na	92.8	74.4	74.0	82.8	69.2
U14	-4.96	5.62	4.54	4.53	4.45	4.55/4.09	101.7	93.3	74.9	71.5	81.4	63.4
U15	-4.11	5.78	4.63	4.62	4.49	4.56/4.16	103.3	92.4	75.4	72.5	81.6	64.6
G16	-3.88	5.80	4.59	4.33	4.54	4.47/4.20	na	92.9	74.7	72.9	82.2	66.4
U17	-4.82	5.46	4.37	4.52	4.44	4.46/4.10	102.3	93.1	74.6	71.3	81.4	63.4
C18	-4.40	5.60	4.39	4.48	4.46	4.55/4.09	96.7	93.3	75.0	71.4	81.2	63.6
C19	-4.40	5.50	4.30	4.48	4.42	4.57/4.08	96.9	93.6	75.1	71.8	81.4	64.3
3' C20	-4.21	5.74	4.04	4.20	4.18	4.50/4.04	97.4	92.4	76.8	68.8	82.8	64.4

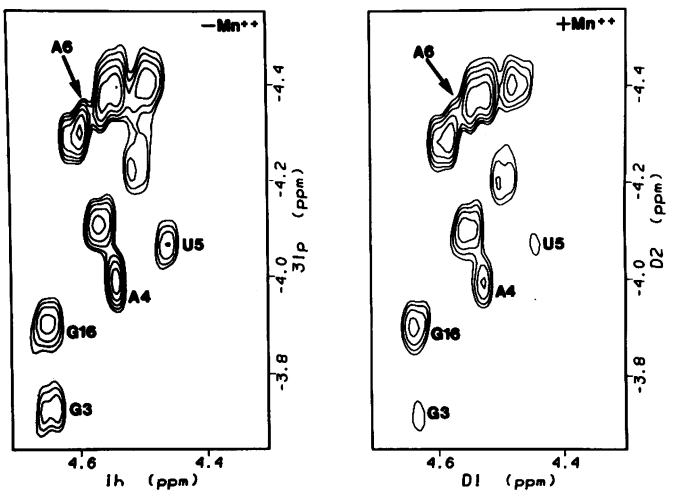

Figure 7. Comparison of ^{1}H–^{31}P correlated hetero-TOCSY spectra identifying the phosphorous resonances (G3 and U5) located in close proximity to the metal binding sites.

ribozymes (Fig. 1). The first site is located in the vicinity of the U•G base pair that defines the site of cleavage, the second at three successive guanines at the 5′-end of the stem. The comparison of the wild type P1 sequence with four different mutants confims that the two binding sites have comparable affinity and are independent. When the U•G wobble pair was replaced by C•G or U•A pairs, the paramagnetic broadening at the position corresponding to the U•G base pair was not observed, but broadening at the GGG sequence at the end of the stem was conserved. In contrast, mutations at the bottom of the stem did not significantly affect manganese binding at the U•G wobble pair.

The Mn^{2+} binding site at the base of the stem is similar to several metal binding sites for Zn^{2+} in tRNA (44). The strong broadening of resonances neighboring the N7 position of the three guanines at the 5′-end of the stem indicate that the manganese is in close proximity to the N7 imidazole nitrogens of the three consecutive guanines. The decreased ligand specificity of Mn^{2+} and Zn^{2+} in comparison with Mg^{2+} suggests that the binding site at the base of the stem is more specific for transition metals than for Mg^{2+}. The second manganese binding site is located in the major groove of the U•G base pair (Fig. 6). This site differs from that located at the base of the stem, since resonances close to the guanine N7 are not significantly broadened in the presence of Mn^{2+}. The investigation of Mn^{2+} binding to a G•U base pair within the RNA acceptor stems (14,15) shows several similarities with our results: although the tRNA investigation was limited to imino resonances, effects similar to those we observed were found at the G•U base pair and at neighboring purine–purine steps (14,15).

Although it is difficult to estimate the affinity of Mn^{2+} for this RNA oligonucleotide, the metal is kinetically in fast exchange on the NMR time scale: this implies a dissociation constant $K_d \geq 10^{-4}$ M, assuming that k_{on} is diffusion limited. Given the relatively weak affinity, it seems that the manganese is not covalently coordinated to the RNA. Rather, the NMR data identify regions of the RNA major groove which constitute sites of high occupancy for the divalent ions. This mode of binding is distinct from both the highly specific interactions observed between Mg^{2+} and certain RNA tertiary structures (41,45), but also from the weak electrostatic binding to charged polyelectrolytes (41) that leads to non-selective broadening of all NMR resonances at high concentrations of manganese (Fig. 2a).

The results of this work reveal the existence of a manganese binding site in the major groove of the U•G base pair that defines the upstream site of cleavage for all group I self-splicing introns.

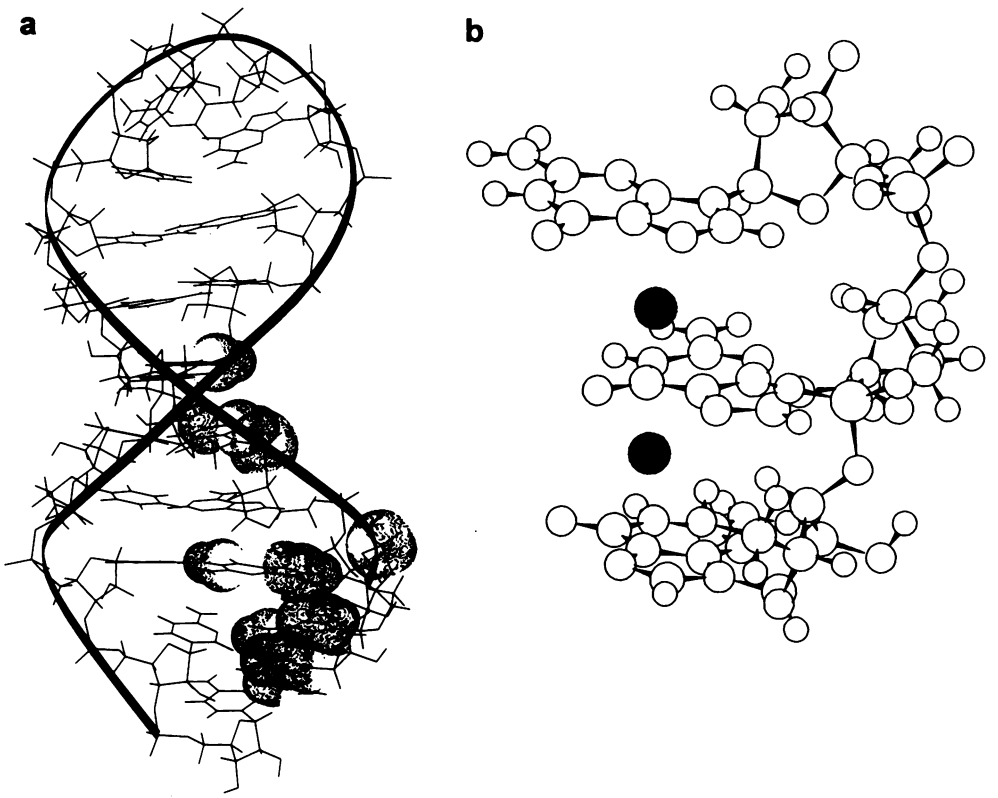

Figure 8. (**a**) View of the structure of the P1 helix, with Van der Waals surfaces identifying the nuclei in closest contact with the Mn^{2+} ions. (**b**) This view of the consecutive guanines at the bottom of the stem highlights the similarities with the divalent ion binding sites identified crystallographycally for tRNA.

Based on the different ability of sulfur and oxygen to coordinate magnesium and manganese, it has been shown that the transition state is stabilized by metal binding to the O3′ of the uracil base and/or to one of the non-bridging oxygens of the phosphate group between the uracil and the following adenine (3,12). In our sequence, this location corresponds to the U5 and A6 bases (Fig. 1), and although no effect on the A6 phosphorus resonance was detected, a clear Mn^{2+} effect was detected on U5 H3′, indicating that the divalent metal ion and the uracil O3′ are close. Metal ion coordination or a metal bridge may be used by the ribozyme to recognize the precise location of the U•G base pair (4). Metal ions promote RNA recognition by certain RNA-binding proteins (41), and the ability of G•U base pairs to coordinate metals may also be important for $tRNA^{Ala}$ recognition by its cognate amino acyl tRNA synthetase (14,15).

From the methodological point of view, we have presented a powerful extension of a technique to identify metal binding sites in RNA. RNA enzymes are metalloenzymes (1), and the elucidation of the structure and thermodynamic features of metal binding to RNA is essential to understand ribozyme function. The methodology described in this work is based on the observation of paramagnetic broadening induced by Mn^{2+} ions upon binding to RNA. This methodology had in the past been applied only to the study of the paramagentic broadening of imino proton resonances. The present results show that the location of metal binding sites in RNA can be identified to a considerable degree of precision, if all accessible NMR resonances are monitored after complete spectral assignments have been obtained. Applications of this approach to small ribozymes such as the hammerhead (46) or the leadzyme (47,48) could reveal structural and thermodynamic aspects of the metal–RNA interactions that are essential for RNA folding and catalysis (49,50).

CONCLUSIONS

The paramagnetic broadening of the NMR resonances induced by manganese binding has been used to characterize the interaction between Mn^{2+} and an oligonucleotide representing the substrate for group I intron ribozymes. Metal binding occurs in the major groove of three consecutive G residues, and at the U•G base pair that defines the 5′-splice site in all group I self-splicing introns. Binding occurs in the fast exchange NMR time scale, corresponding to a dissociation constant $K_d \geq 10^{-4}$ M. Thus, the interaction is not a static coordination complex between the RNA and the metal. Rather, the NMR data identify part of the helical major groove as a region of high occupancy for the divalent ions. The binding site in the vicinity of consecutive Gs appears to be very similar to crystallographically identified Zn^{2+} binding sites in tRNA where the metal was coordinated via its hydration shell to the N7 of successive guanines. In addition to tertiary contacts involving several critical 2′-hydroxyl groups, metal ion coordination or a metal bridge may be used by group I introns in recognizing the absolutely conserved U•G base pair that defines the 5′ splice site. Extension of the paramagnetic broadening technique to monitor not only imino resonances, as often done in the past, but all assignable resonances, allows a precise identification of the metal binding sites. The application of this technique to RNA enzymes would be of great interest to understand how

RNA structure defines metal binding sites for catalysis and folding.

ACKNOWLEDGEMENTS

It is a pleasure to thank Dr Fareed Aboul-ela and Mr Charles Gubser for their help and suggestions. FA was supported by fellowships from the École Normale Supérieure, Paris, France and by the Medical Research Council.

REFERENCES

1 A. M. Pyle (1993) *Science* **261**, 709–714.
2 H. W. Pley, K. M. Flaherty, D. B. McKay (1994) *Nature* **372**, 68–74.
3 J. A. Piccirilli, J. S. Vyle, M. H. Caruthers, T. R. Cech (1993) *Nature* **361**, 85–88.
4 E. L. Christian, M. Yarus (1993) *Biochemistry* **32**, 4475–4480.
5 T. R. Cech (1990) *Annu. Rev. Biochem.* **59**, 543–568.
6 F. Michel, E. Westhof (1990) *J. Mol. Biol.* **216**, 585–610.
7 A. M. Pyle, F. L. Murphy, T. R. Cech (1992) *Nature* **358**, 123–128.
8 S. A. Strobel, T. R. Cech (1993) *Biochemistry* **32**, 13593–13604.
9 S. A. Strobel, T. R. Cech (1994) *Nature Struct. Biol.* **1**, 13–17.
10 J. A. Doudna, B. P. Cormack, J. W. Szostak (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7402–7406.
11 T. A. Steitz, J. A. Steitz (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6498–6502.
12 T. Uchimaru, M. Uebayasi, K. Tanabe, K. Taira (1993) *FASEB J.* **7**, 137–142.
13 R. E. Hurd, E. Azhderian, B. R. Reid (1979) *Biochemistry* **18**, 4012–4017.
14 S. Limmer, H.-P. Hofmann, G. Ott, M. Sprinzl (1993) *Proc. Nat. Acad. Sci. USA* **90**, 6199–6202.
15 G. Ott, L. Arnold, S. Limmer (1993) *Nucleic Acids Res.* **21**, 5859–5864.
16 C. Tuerk, *et al.* (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1364–1368.
17 J. F. Milligan, D. R. Groebe, G. W. Witherell, O. C. Uhlenbeck (1987) *Nucleic Acids Res.* **15**, 8783–8789.
18 J. D. Puglisi, I. Tinoco Jr (1990) *Methods Enzymol.* **180**, 304–325.
19 R. T. Batey, M. Inada, E. Kujawinski, J. D. Puglisi, J. R. Williamson (1992) *Nucleic Acids Res.* **20**, 4515–4523.
20 E. P. Nickonowicz, *et al.* (1992) *Nucleic Acids Res.* **20**, 4507–4513.
21 J. V. Hines, S. M. Landry, G. Varani, I. Tinoco Jr (1994) *J. Am. Chem. Soc.* **116**, 5823–5831.
22 G. Varani, C. Cheong, I. Tinoco Jr. (1991) *Biochemistry* **30**, 3280–3289.
23 D. Marion, K. Wüthrich (1983) *Biochem. Biophy. Res. Comm.* **113**, 967–974.
24 P. Plateau, M. Gueron (1982) *J. Am. Chem. Soc.* **104**, 7310–7311.
25 A. J. Shaka, C. J. Lee, A. Pines (1988) *J. Magnetic Resonance* **77**, 274–293.
26 V. Sklénar, H. Miyashiro, G. Zon, A. Bax (1986) *FEBS Lett.* **208**, 94–98.
27 G. W. Kellogg (1992) *J. Magnetic Resonance* **98**, 176–182.
28 G. W. Kellogg, B. I. Schweitzer (1993) *J. Biomolecular NMR* **3**, 577–595.
29 G. Bodenhausen, D. J. Ruben (1980) *Chem. Phy. Lett.* **69**, 185–189.
30 A. M. Gronenborn, A. Bax, P. T. Wingfield, G. M. Clore (1989) *FEBS Lett.* **243**, 93–98.
31 A. Bax, M. Ikura, L. E. Kay, D. A. Torchia, R. Tschudin (1990) *J. Magnetic Resonance* **86**, 304–318.
32 E. P. Nikonowicz, A. Pardi (1993) *J. Mol. Biol.* **232**, 1141–1156.
33 V. Sklénar, R. D. Peterson, M. R. Rejante, J. Feigon (1994) *J. Biomolecular NMR* **4**, 117–122.
34 A. J. Shaka, P. Barker, R. Freeman (1985) *J. Magnetic Resonance* **64**, 547–552.
35 A. Bax, G. M. Clore, A. M. Gronenborn (1990) *J. Magnetic Resonance* **88**, 425–431.
36 A. Pardi, E. P. Nikonowicz (1992) *J. Am. Chem. Soc.* **114**, 9202–9203.
37 E. R. P. Zuiderweg, L. P. McIntosh, F. W. Dahlquist, S. W. Fesik (1990) *J. Magnetic Resonance* **86**, 210–216.
38 S. J. Archer, D. M. Baldisseri, D. A. Torchia (1992) *J. Magnetic Resonance* **97**, 602–606.
39 J. P. Marino, *et al.* (1994) *J. Am. Chem. Soc.* **116**, 6472–6473.
40 G. Varani, I. Tinoco Jr (1991) *Q. Rev. Biophys.* **24**, 479–532.
41 L. G. Laing, T. C. Gluick, D. E. Draper (1994) *J. Mol. Biol.* **237**, 577–587.
42 A. Jack, J. E. Ladner, D. Rhodes, R. S. Brown, A. Klug (1977) *J. Mol Biol* **111**, 315–318.
43 B. Hingerty, R. S. Brown, A. Jack (1978) *J. Mol. Biol* **124**, 523–534.
44 J. R. Rubin, J. Wang, M. Sundaralingam (1983) *Biochim. Biophys. Acta* **756**, 111–118.
45 Y. Chen, H. Sierzputowska-Gracz, R. Guenther, K. Everett, P. F. Agris (1993) *Biochemistry* **32**, 10249–10253.
46 O. C. Uhlenbeck (1987) *Nature* **328**, 596–600.
47 T. Pan, O. C. Uhlenbeck (1992) *Nature* **358**, 560–563.
48 T. Pan, B. Dichtl, O. C. Uhlenbeck (1994) *Biochemistry* **33**, 9561–9565.
49 S. C. Dahm , O. C. Uhlenbeck (1991) *Biochemistry* **30**, 9464–9469.
50 S. A. C. Dahm, W. B. Derrick, O. C. Uhlenbeck (1993) *Biochemistry* **32**, 13040–13045.